

Background

With the increasing kidney shortage crisis present in the United States that is further complicated with the occurrence of the COVID-19 pandemic, immediate solutions are required to both reduce kidney discard and increase the success rate of the transplants that do occur. Solutions accomplishing both goals may be found in the massively growing realm of artificial intelligence (AI) and its subfield machine learning (ML). While current clinical implementations are limited, significant efforts are being made across all of nephrology and related fields to utilize the powerful tool that is AI to improve patient care and outcomes across the board¹. An area with especially high opportunity for the application of AI is renal transplant pathology. In this project, we apply AI to classify morphologic features of deceased donor renal biopsies in order to predict the likelihood of DGF associated with the biopsied kidney.

Definitions and Goals

A well-known indicator of long-term graft function in both deceased and living donor kidney transplants is delayed graft function (DGF)². The United Network for Organ Sharing (UNOS) defines DGF as the need for at least one dialysis treatment in the first week after transplantation³. A trained expert pathologist is able to analyze biopsies of donor kidneys and accurately predict the viability of the graft in isolation from clinical variables. This is done through recording features similar to those in Table 1, which measures histopathologic features shown to affect the viability of the graft.

Histopathologic Finding	Report format
Global Sclerosis	Percent Sclerosed
Segmental Sclerosis	Yes/No
Interstitial Fibrosis	Score: 0-3
Tubular Atrophy	Score: 0-3
Arterial Fibrosis	Score: 0-3
Arteriolar Hyalinosis	Yes/No

Table 1: Example pathologist donor biopsy report sheet. The end goal of this AI is to emulate a trained pathologist's ability to analyze tissue enough to fill out this form.

The availability of pathologists trained to the level needed to accurately assess a donor renal biopsy at any hour of the day is relatively low, so the availability of an AI-based software to accurately make this assessment would significantly improve the availability of consistent and accurate predictions of DGF and graft function in general. The goal of both part I and II of this project is to do exactly that, with the difference in the two parts being the stain used for the visualization of the pathology. Part I is done with the Periodic Acid-Schiff stain and part II is done with the Hematoxylin and Eosin stain.

Previous Work and Staining

Since 2016, many papers have been published in AI Renal Pathology. Most of these papers focused primarily on segmentation of glomeruli only, with few progressing to more detailed than that and even fewer working on applications in transplant. Another important aspect of AI analysis of pathology is the stain used for visualization. The generally accepted best stain for AI is Periodic Acid-Schiff (PAS) stain for its well defined borders and color differentiation between different features. The PAS stain is the most used stain in the literature for these reasons. Almost all renal biopsies are routinely frozen and stained with Hematoxylin and Eosin (H&E), which offers fewer differentiation between features, and the freezing process can often destroy or alter the features that are visible (Figure 1).

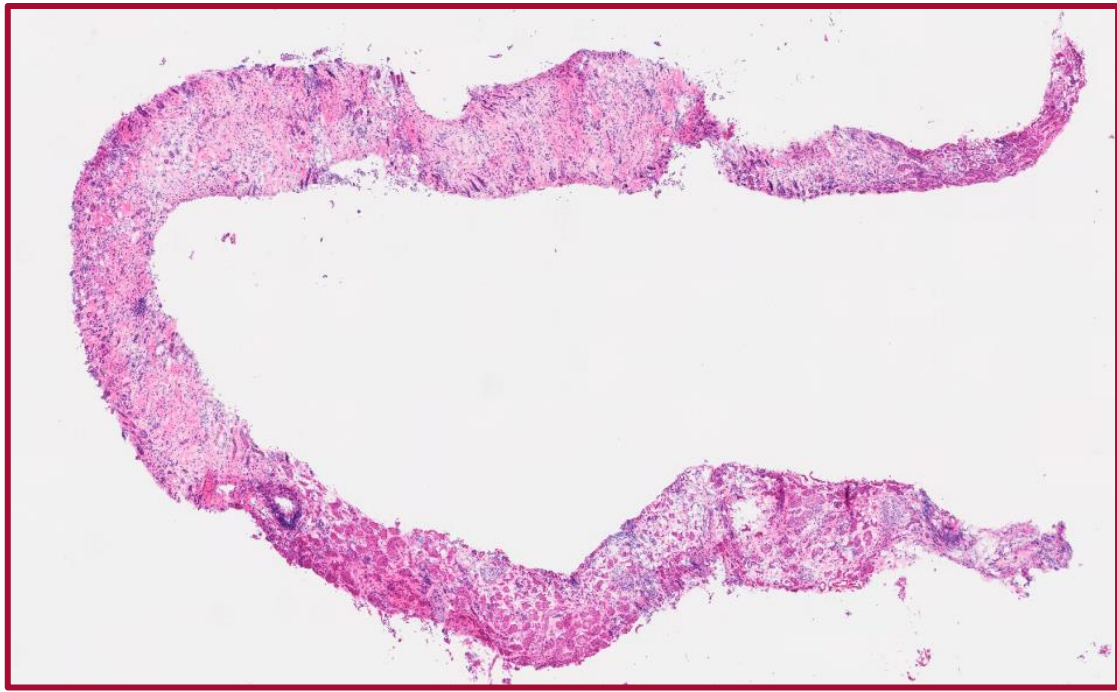


Figure 1: Example frozen H&E stained renal biopsy

Results

I. Advanced Segmentation with PAS stain

The algorithm developed using the PAS stain was able to very accurately identify many different morphologic features of the renal biopsies.

Glomeruli: Accurate identification and count within ± 1 from the true number

Globally Sclerotic Glomeruli: Accurate identification and count within ± 1 from the true number

Segmental Sclerosis: Can accurately identify binary presence or absence, but can not quantify with clinical level accuracy

Interstitial Fibrosis: Can accurately quantify on a 0-3 scale but can be thrown off by certain underlying pathologies

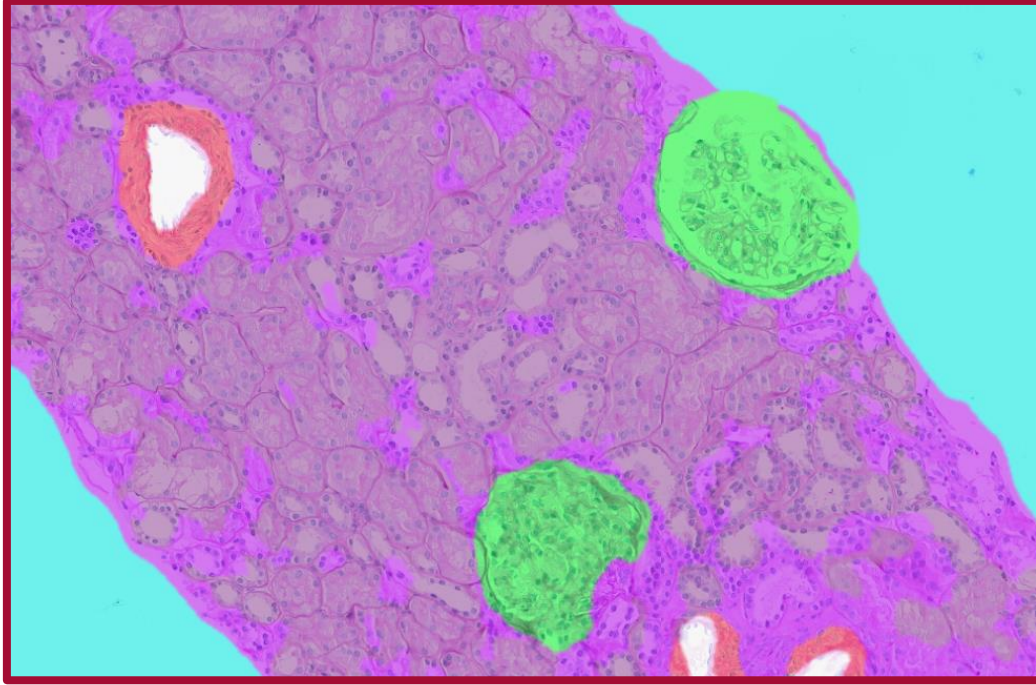


Figure 5: Interstitial fibrosis, tubular atrophy, and arterial fibrosis segmentation masks example PAS stained slide



Figure 4: Glomeruli segmentation masks example PAS stained slide

Tubular Atrophy: Can accurately quantify on a 0-3 scale but can be thrown off by certain underlying pathologies

Arterial Fibrosis: Identifies by seeing increased intimal to luminal area ratio to give a scaled score from 0-3. Accuracy data pending but preliminarily optimistic.

Arteriolar Hyalinosis: Can identify binary presence or absence with high accuracy

Overall, this methodology and procedure for identification shows extreme promise for application to part (II) in identification on H&E

II. H&E Autofluorescence

The second part of this project is to investigate the ability of H&E to be a valid stain to do everything that we were able to accomplish with the PAS stain. It is hypothesized that this can be done more accurately with scanning the slides using a fluorescence microscope to view the autofluorescence intrinsic to the H&E stain. This trial compares the ability of AI networks to segment H&E stained samples with different methods of visualization.

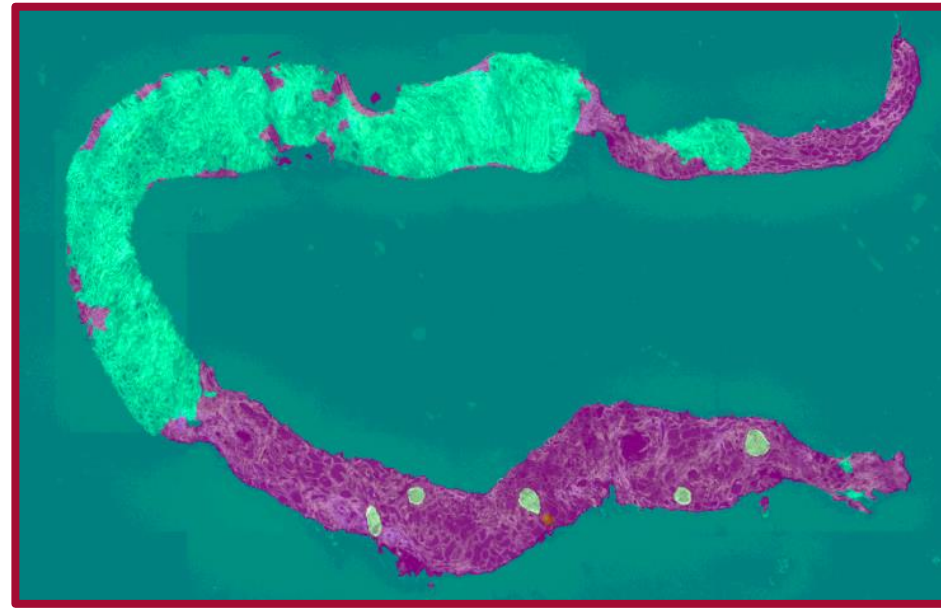


Figure 6: Example frozen H&E biopsy scanned with fluorescence microscopy with segmentation masks

Classifier	Visualization Method		
	Bright field H&E	1 Channel IF H&E	5 Channel IF H&E
Background	0.970	0.973	0.994
Renal Tissue	0.980	0.987	0.989
Normal Glom.	0.906	0.920	0.926
Globally Sclerotic	0.552	0.897	0.892
Overall	0.852	0.945	0.950

Table 2: F1 scores of the different visualization methods for each of the different segmentation classes. The closer the F1 score is to 1, the higher the accuracy

The segmentation accuracy is measured using the F1 score, which is a weighted average of the precision and recall of the network such that the closer the number is to 1, the higher the accuracy of identification. In this study it was found that the networks trained and tested on fluorescence microscopy scanned samples performed significantly better across the board. This was especially apparent with the identification of globally sclerotic glomeruli, with which the bright field visualization method struggled significantly. It should also be noted that although the five-channel scan did perform better than the single channel in most categories, the results were comparable enough that the single channel has high potential as a cost – effective alternative to the five-channel option.

Conclusions and Next Steps

This work has shown that artificial intelligence through the HALO AI software is currently in a state such that it can be applied to accurately segment a variety of renal morphologic features, and it has shown promise in the ability to do so even with the H&E stain through the use of autofluorescence. The next steps of this project are to further develop the H&E AI to the quality level of the one done on PAS through both giving more training samples and using expert pathologists to create the training regions for the AI. Once the network is developed to the point of clinical level accuracy of segmentation, it will then be used along with a parallel project to be able to automatically predict the chance of DGF occurring in any given renal transplant using just the biopsy and a few donor and recipient characteristics.

The end goal is to have an affordable kit that can be bought by any hospital in the country that supports organ transplantation to be able to scan their biopsy and donor/recipient characteristics and immediately get information on the potential viability of the graft in their patient.

References

1. Checucci E, et. al. *Minerva Urol Nefrol.* 2020;72(1):49-57.
2. Senel FM, et. al. *Ren Fail.* 1998;20(4):589-595.
3. Kim, W. R., et. al. *Am J Transplant*, 15, 1-28.
4. Meng H, et. al. *J Vis Exp.* 2014;(89):51586.
5. Jeldres C., et. al. *Can Urol Assoc J.* 2009;3(5):377-382.
6. Chapal M, et. al. *Kidney Int.* 2014;86(6):1130-1139.
7. Zaza G, et. al. *Intern Emerg Med.* 2015;10(2):135-141.
8. Karpinski J, et. al. *Transplantation.* 1999;67(8):1162-1167.
9. Munivenkatappa RB, et. al. *Am J Transplant.* 2008;8(11):2316-2324.
10. Chen K, et. al. *Transplant Proc.* 2021;53(3):906-912.
11. Salvi M, et. al. *Comput Med Imaging Graph.* 2021;90:101930.

Methods

HALO AI is a collection of train-by-example classification and segmentation tools underpinned by advanced deep learning neural network algorithms. HALO



includes multiple pre-built AI networks, including their DenseNet V2 network - a pre-built densenet network with utilizes exceptionally dense and convoluted connections between layers, which was used in this project. The DenseNet V2's increased complexity over alternatives allows for the picking up of the most minute details in the tissue, given the drawback of longer training times and more training regions required for success.

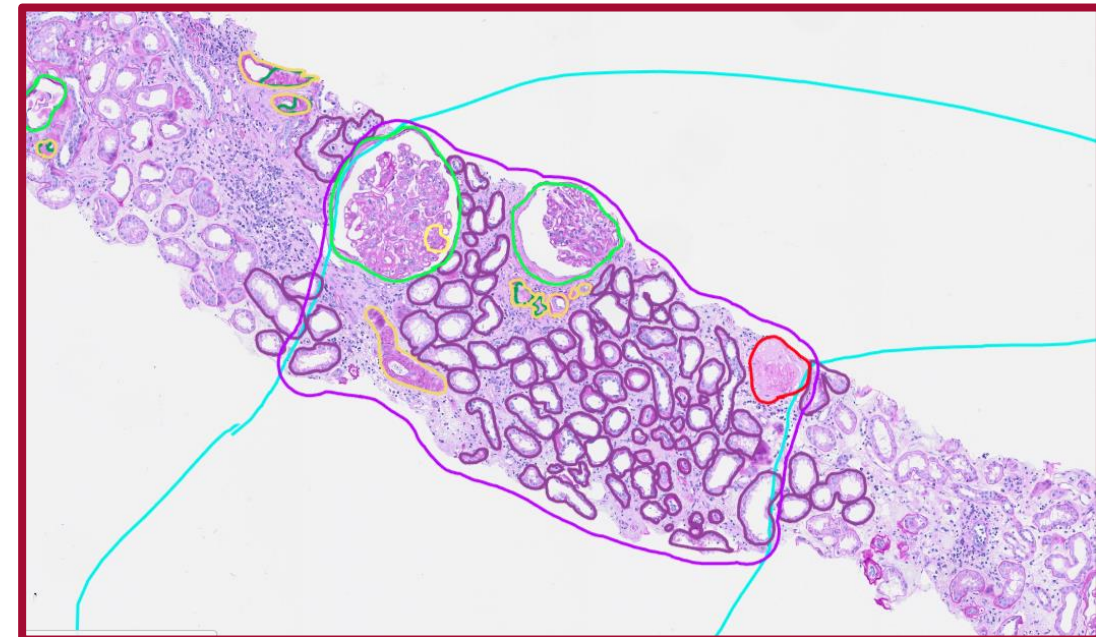


Figure 2: Example of annotation of many different classes on a PAS slide

I. Advanced Segmentation with PAS stain

HALO was used to train semantic segmentation AI using the DenseNet V2 network on PAS-stained images from a data set of donor biopsies for all fields described in Table 1 – background, tissue, globally sclerotic glomerulus, segmentally sclerotic glomerulus, Interstitial Fibrosis, Tubular Atrophy, Arterial Fibrosis, and Arteriolar Hyalinosis. Images were manually annotated similarly to that shown in Figure 2 for each class and trained over 10,000 iterations. The algorithm output was then analyzed using custom written scripts to get numerical values for those that are numerical in Table 1 and binary for those that are binary.

II. H&E Autofluorescence

In order to combat AI's inherent difficulties with frozen H&E stain, we investigated using the autofluorescence intrinsic to the H&E stain to analyze the biopsies. HALO was used with the same DenseNet V2 network to train semantic segmentation AI using the on three different imaging modalities - bright field, single channel fluorescence, and five-channel fluorescence - over the same regions of the H&E stained samples from a UAMS data set of donor kidney biopsies. Images were manually annotated for the segmentation categories background, tissue, non-sclerotic glomerulus, and globally sclerotic glomerulus in each imaging modality separately. This was done to intentionally include the differences in difficulty of annotating between the imaging modalities as a variable determining the performance of the network. Each network was then trained over equal sample size and 2500 training iterations, while concurrently being validated over the same validation sample set. The validation metric used was the F1-score, a metric developed by HALO to evaluate network performance such that the closer the score is to one, the better the performance.

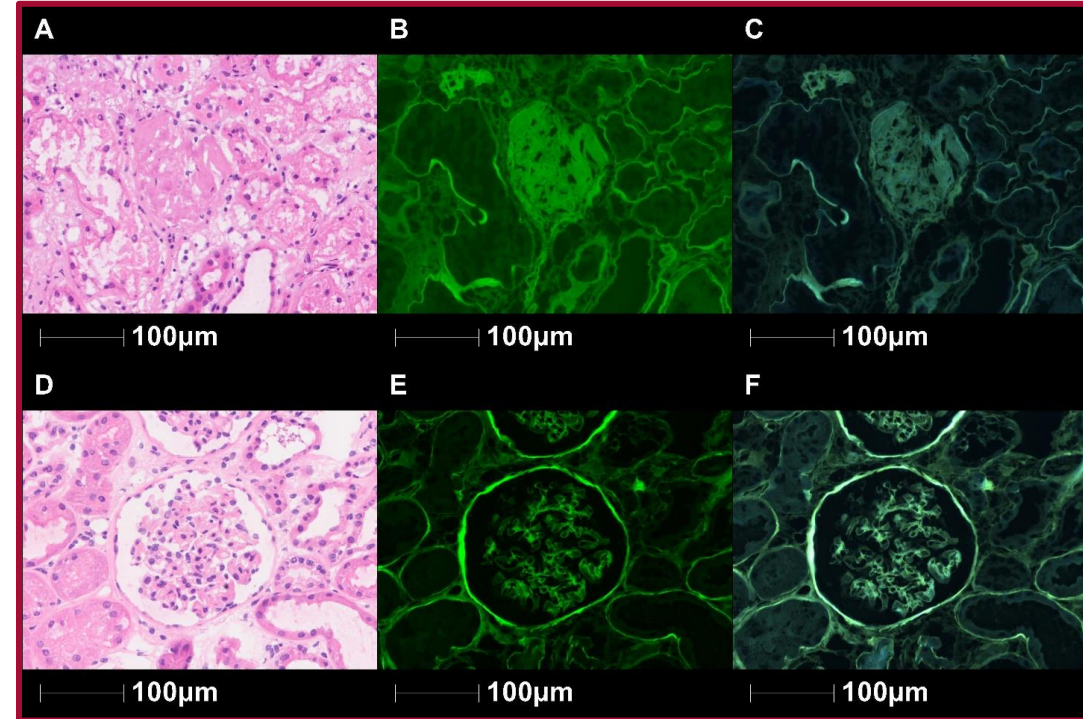


Figure 3: Image of the same globally sclerotic glomerulus under (A) bright field H&E, (B) single channel fluorescence, and (C) five channel fluorescence and image of a normal glomerulus under (D) bright field H&E, (E) single channel fluorescence, and (F) five channel fluorescence.

Acknowledgments

This work was supported in part by the 2022 Urology Care Foundation Summer Medical Student Fellowship Program and the Arkansas Urologic Society. A special thanks to UAMS and Arkana Laboratories for their support in this project.

Contact: dcampbell5@uams.edu